

Developing the Next Generation of Physics Assessments

James T. Lavery,¹ Melanie M. Cooper,^{2,1} and Marcos D. Caballero^{3,1}

¹*CREATE for STEM Institute, Michigan State University, East Lansing, MI 48824, USA*

²*Department of Chemistry, Michigan State University, East Lansing, MI 48824, USA*

³*Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA*

Science education at all levels is currently undergoing dramatic changes to its curricula and developing assessments for these new curricula is paramount. We have used the basis of many of these new changes (scientific practices, crosscutting concepts, and core ideas) to develop sets of criteria that can be used to guide assessment development for this new curriculum. We present a case study that uses the criteria we have developed to revise a traditional physics assessment item into an assessment item that is more aligned with the goals of current transformation efforts. Assessment items developed using this criteria can be used to assess student learning of both the concepts and process of science.

PACS numbers: 01.40.G-, 1.40.Di, 1.40.Fk

Keywords: assessment, three-dimensional learning, practices, core ideas, STEM

I. INTRODUCTION

The science education curriculum is currently undergoing a significant overhaul in the ways students are taught and learning is assessed. In 2012, the National Research Council released *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (the *Framework*), which described, in detail, three dimensions of science and engineering learning on which to base a curriculum[1]. The three dimensions are briefly described here, followed by examples from the document:

Scientific Practices These eight practices can be thought of as the disaggregated components of the process of science. These practices engage students in science by putting their knowledge to use to model, predict, and explain phenomena. *Examples: Planning and Carrying Out Investigations; Constructing Explanations*

Crosscutting Concepts These seven concepts bridge the boundaries between the disciplines of the physical, biological, and geological sciences. These “ways of thinking” are used by each of the disciplines and can be leveraged to help students make connections across them. *Examples: Patterns; Cause and Effect; Mechanism and Explanation*

Disciplinary Core Ideas These are the foundationally important concepts that are fundamental to members of a scientific discipline. In order to qualify as a disciplinary core idea, the concept must be essential to the study of the discipline, be required to explain a wide range of phenomena, and provide a way to generate new ideas and predictions. *Examples: Matter and its Interactions; Energy*

The *Framework* emphasizes that it is vital that all three dimensions are blended into nearly every aspect of students’ learning opportunities (which we call “three-dimensional learning”).

A. Bringing Three-Dimensional Learning to College Classrooms

While the *Framework* was written for the K-12 community, its conclusions (which were informed by both science education and discipline-based education research) can and should be applied to introductory college science courses as well[2]. After all, college instructors certainly would like to help guide their students to, for example, be able to construct explanations (scientific practice) for systems (crosscutting concept) that utilize the concept of energy (disciplinary core idea).

With the adoption of the Next Generation Science Standards (which are based on the *Framework*) by more and more states, the K-12 education system is currently heading in the direction of implementing three-dimensional learning at the national scale[3]. In our view, universities and colleges can and should react to this for two reasons. First, as time passes, more students entering college will have experienced three-dimensional learning. If universities and colleges aim to capitalize on this new form of science education, we must start adapting and adopting now. Second, with the significant effort needed to incorporate the new standards, K-12 teachers will need to be trained not only in the knowledge of science, but also its process. These teachers are often required to take introductory college science courses, which means we have an opportunity (duty?) to help develop the next generation of science teachers.

B. Assessing Three-Dimensional Learning

At the college level, assessments have helped drive transformation efforts and measure their impact in the past. In fact, much of the early history of discipline-based education research focused on the development of assessments and adjusting curricula to help students improve on those assessments[4, 5]. These ‘concept inventories’ are often suggested and used as catalysts or assessments

for efforts to transform teaching and learning[6]. Thus, being able to assess three-dimensional learning should be a high priority for modifying instruction. After all, if we do not assess what is important, then what is assessed becomes important. Additionally, as assessments are a part of students' learning opportunities, they often drive the 'hidden curriculum', which has a significant effect on student learning[7].

The remainder of this paper will introduce work that endeavours to help instructors (and other assessment authors) develop assessment items (e.g., exam, homework, and clicker questions) that align with three-dimensional learning. Given the format for this paper, we will present only a single case study, specifically for a college level physics course. To do this, first we introduce criteria from a newly developed instrument, then use these criteria to analyze a typical introductory physics problem, and finally modify that problem to better align with the three dimensions.

II. DEVELOPING THREE-DIMENSIONAL ASSESSMENTS

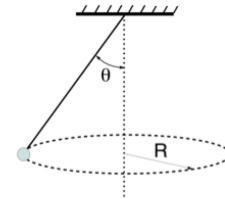
A. Assessment Item Criteria

As part of a larger transformation project at Michigan State University, we have been developing the Three-Dimensional Learning Assessment Protocol (3D-LAP), a soon-to-be released research instrument that can be used to both design and characterize assessment items. The principle components of the 3D-LAP are sets of criteria that have been developed for each scientific practice, crosscutting concept, and core idea. These criteria can be used by assessment authors to help modify or generate items that align with the three dimensions of the *Framework*. Items that meet these criteria align with the scientific practices, crosscutting concepts, and core ideas as well as give instructors more explicit feedback about what students are capable of doing.

The 3D-LAP uses the list of scientific practices and the list of crosscutting concepts from the *Framework*. Because there is no national document that has a list of core ideas for college level physics, we are left to invent our own. At Michigan State University, faculty in the Physics and Astronomy department have discussed and generated the set of core ideas for our courses[8].

There is a fundamental difference between the three dimensions that results in different formats for the criteria. Because the scientific practices are about the process of science, they require that certain actions be performed ('verbs'). As a result, the criteria for the scientific practices all focus on the activities the assessment item asks the student to engage in. Unlike the scientific practices, crosscutting concepts and core ideas are constructs that represent scientific knowledge ('nouns'). As such, the criteria for them include subconstructs that must either be indicated in the question, or asked of the student (in our

1. A mass of 2.89 kg is suspended from a 1.07 m long string. It revolves in a horizontal circle as shown in the figure.



If the string makes an angle of $\theta=27.5^\circ$ with the vertical, then what is the net force acting on the mass? (in N)

- A. 10.8 B. 12.6 C. 14.8 D. 17.3

FIG. 1: A traditional introductory physics assessment item that aligns with one of the three dimensions in the *Framework*.

example this means it must be part of the question stem or the correct answer choice.) Sample criteria from the 3D-LAP can be found in the first column of Table I.

B. A Traditional Assessment Item

Figure 1 shows a typical introductory physics question for which an instructor might hope students would build a free body diagram, use Newton's second law to generate a mathematical representation, and solve those equations for a particular value. Unfortunately, this item does not, in fact, give explicit evidence that a student is doing any of those things, as choosing the correct or incorrect solution (in this case) does not help us understand how a student arrived at that particular answer.

In order to accomplish this and, in the process, align with each of the dimensions of the *Framework*, we must first decide which scientific practice, crosscutting concept, and core idea this question already aligns with most closely (or, alternatively, which ones we would like it to align with). For this example, we have chosen:

Scientific Practice Developing and using models
Crosscutting Concept Systems and system models
Disciplinary Core Idea Interactions can cause changes in motion

The central column of Table I contains our analysis of the question in Figure 1, including evidence for why it does or does not meet each particular piece of the criteria.

In this particular case, the question in Figure 1 is most closely connected to our (locally determined) core idea of *Interactions can cause changes in motion*[8]. While one might argue that this is not the proper core idea, or that they would use a different one, we strongly suspect that every such list would include the idea of forces, interactions, and their effects on the motions of objects. For this reason, we argue that this traditional item already contains a core idea and do not include the criteria or analysis in Table 1 due to space considerations.

TABLE I: An analysis of the assessments in this article using the criteria we have developed for the 3D-LAP instrument. An assessment item must meet all of the criteria for a dimension in order for it to be considered ‘Aligned’ with that dimension.

3D-LAP criteria for aligning with three-dimensional learning	Elements of Traditional Item (Figure 1) that meet the criteria	Elements of Revised Item (Figure 2) that meet the criteria
Developing and using models:	Unaligned	Aligned
1. Question gives an event, observation, or phenomenon for the student to explain or make a prediction about.	✗ 1. Question describes a system that is already idealized.	✓ 1. Question stem describes “the faster the ball moves, the larger the angle the rope makes with the pole.”
2. Question gives student a representation or asks student to select a representation.	✓ 2. Question includes a visual representation of the situation.	✓ 2. Question 1 asks “Which of the following free body diagrams”
3. Question asks student to select an explanation for or prediction about the event, observation, or phenomenon using the representation.	✓ 3. Question asks “what is the net force acting on the mass?”	✓ 3. Answer 3a includes “As the speed increases, the angle must increase.”
4. Question asks student to select the reasoning that links the representation to their explanation or prediction.	✗ 4. Question does not ask for any evidence of reasoning.	✓ 4. Answer 1b (free body diagram), answer 2a (equation), and answer 3a includes “in order for the net force to maintain the ball’s circular trajectory.”
System and system models:	Unaligned	Aligned
1. The components of the system.	✓ 1. Question asks “acting on the mass?”	✓ 1. Question 1 asks “Which of the following free body diagrams would apply to the ball”
2. The interactions between those components.	✓ 2. <i>Does not apply, since there is only one object in the system.</i>	✓ 2. <i>Does not apply, since there is only one object in the system.</i>
3. The interactions of those components with the surroundings.	✗ 3. Question and answer do not include evidence of the interactions.	✓ 3. Answer 1b represents the interactions with the environment.

C. Revising an Assessment Item

In order to obtain more explicit evidence of student reasoning, the revised item was developed into a cluster of three related questions. Figure 2 contains the revised version of the assessment item, which we claim aligns with the three dimensions. The analysis of how it aligns with the chosen scientific practice and crosscutting concept can be found in the right most column of Table I.

For the scientific practice, the original question is not meeting two of the criteria. First, the situation described is already idealized (moving in a perfect circle with no friction). One part of the modification must be to place the question in a real world setting (which might then be modeled as movement in a circle with no friction). The second missed criterion requires evidence about the reasoning the student is using to answer the question. Because this is a selected response setting, multiple options for the reasoning students might give must be included as part of the set of answers.

As for the crosscutting concept, the original question does not include any information about the interactions between the system and its surroundings. That information must be provided in some way or another. For this particular revision, we have chosen to have the student select the appropriate free-body diagram, which contains this information.

Ultimately, we claim (based on the analysis found in Table I) that the revised cluster of items in Figure 2 is more capable of assessing a student’s ability to engage


with a scientific practice, crosscutting concept, and core idea than the traditional item in Figure 1.

III. DISCUSSION


Comparing the traditional and revised assessment item, one might argue that we have altered the intent of the question because we no longer ask the student to solve numerically for a particular value. However, it is important to note that while such a question is not necessary to meet the criteria, it could still be asked. In fact, one could imagine adding the traditional question as question number 4 in the revised cluster and the cluster would still meet all of the criteria to qualify as three-dimensional. This points to an interesting question about assessing three-dimensional learning: What about items that we want to assess students with that do not include all three dimensions? We do not claim that all assessment items students engage with need to align with all three dimensions, but clearly some of them need to if we are going to claim that we are assessing three-dimensional learning.

Assessing how a student authentically engages in scientific practices requires that they generate their own ideas in response to a situation where there are no clear answers. Selected response items generally do not work well for probing student engagement with scientific practices because they often require that one answer choice is clearly correct over the other options, something uncommon in science. For this reason, we maintain that


You are watching two of your friends play tetherball and you notice that the faster the ball moves, the larger the angle the rope makes with the pole. In order to figure out why, you begin by thinking about the forces involved.




- Which of the following free body diagrams would apply to the ball swinging freely around the pole?




a.



b.



c.



d.
- Which of the following equations can be extracted from the correct free body diagram (and information provided) above (T is the tension in the rope; r is the distance between the ball and the pole; v is the speed of the ball)?
 - $T \sin \theta = \frac{m v^2}{r}$
 - $T \cos \theta = \frac{m v^2}{r}$
 - $T \sin \theta = 0$
 - $T \cos \theta = 0$
- Combine the equations you obtained from the free body diagram above and solve for θ . Which of the following explains why the angle increases?
 - As the speed increases, the angle must increase in order for the net force to maintain the ball's circular trajectory.
 - As the ball speeds up, the radius decreases and the angle must increase in order for the net force to maintain the ball's circular trajectory.
 - The faster the ball moves, the larger the net force pushes outward on the ball causing it to increase the angle it makes with the vertical.

FIG. 2: The revised cluster of assessment items modified to align with the three dimensions using the criteria from the 3D-LAP.

only constructed response items can engage students in scientific practices.

However, because many introductory science courses are large, constructed response assessment items are often prohibitively expensive and/or time-consuming to grade. Selected response items, on the other hand, require fewer resources to grade, increasing the potential for both faculty productivity and quick turnaround of feedback. For these reasons, assessment items in selected response format are often necessary for these courses.

For the purposes of this paper, we felt it was important to show the modification of a selected response assessment item without changing it to a constructed response

format, as this is the more difficult challenge. Given the argument above, we do not claim that meeting the criteria in this case engages students in a scientific practice, but that it is as close to aligning with the practice as it can be while remaining a selected response item.

While reviewing assessment items during the development of the criteria for the 3D-LAP, one common theme we discovered was that many traditional assessment items (like the one in Figure 1) meet at least some of the criteria for scientific practices. However, many of the scientific practices criteria require evidence of some kind of reasoning (similar to criterion 4 under *Developing and using models* in Table I) and this particular criterion is almost never met by traditional assessment items.

IV. CONCLUSION

Being able to engage in both the process of science (scientific practices) and the knowledge of science (cross-cutting concepts, core ideas) should be the goal of science education. Research is beginning to show that the framework of three-dimensional learning is a productive way to go about designing curricula and assessments to help students accomplish these goals. For this reason and those described earlier, it is necessary to bring three-dimensional learning to our college level physics courses.

If we are going to include it, we need a way to assess it and the criteria we have developed can help to develop the kinds of assessments necessary to promote student learning. In the near future, we will be publishing the 3D-LAP for others to use, critique, and hopefully improve upon so that we can collectively move to the next generation of physics assessments.

V. ACKNOWLEDGEMENTS

The authors thank the Association of American Universities' STEM Education Initiative for their support, the rest of the DBER community at MSU (especially those involved in the development of the criteria), and PERL@MSU for their feedback on early drafts.

-
- [1] National Research Council, *A Framework for K-12 Science Education* (National Academies Press, 2012).
 - [2] M. M. Cooper et al., *Science* **350**, 281 (2015).
 - [3] Achieve, Inc., *Next Generation Science Standards* (National Academies Press, 2013).
 - [4] D. Hestenes, M. Wells, and G. Swackhamer, *The Physics Teacher* **30** (1992).
 - [5] D. P. Maloney, T. L. O'Kuma, C. J. Hieggelke, and A. Van Heuvelen, *American Journal of Physics* **69**, S12 (2001).
 - [6] National Research Council, *Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering* (National Academies Press, 2012).
 - [7] E. F. Redish, J. M. Saul, and R. N. Steinberg, *American Journal of Physics* **66**, 212 (1998).
 - [8] J. T. Laverty, S. H. Tessmer, M. M. Cooper, and M. D. Caballero, in *Physics Education Research Conference 2014* (Minneapolis, MN, 2014), PER Conference, pp. 147–150.